

BAB I

PENDAHULUAN

1.1 Latar Belakang

Meski CCTV sudah dapat terintegrasi dengan berbagai perangkat, tapi penggunaan informasi yang dihasilkan berupa video masih banyak hanya mengandalkan tenaga manusia. Contohnya, jika terdapat suatu barang yang tertinggal ditempat umum, maka petugas keamanan tersebut harus melihat tiap video hasil rekaman CCTV yang ada untuk memastikan di mana barang tersebut tertinggal. Hal tersebut memerlukan waktu yang relatif panjang dan tentunya itu bukanlah pekerjaan yang mudah (Agarwal et al., 2018).

Sudah banyak penemuan yang berkenaan dengan peningkatan teknologi keamanan, salah satunya yaitu pendeteksian objek yang dilakukan hanya dengan mengandalkan tenaga manusia dan digantikan dengan pendekatan kecerdasan buatan. Teknologi kecerdasan buatan yang dapat mendeteksi sebuah objek merupakan teknologi mendasar dalam *computer vision*. *Computer vision* memungkinkan komputer untuk melihat layaknya apa yang dilihat oleh manusia, seperti menangkap, memproses, dan menganalisis. *Computer vision* dapat mengekstraksi informasi dari sebuah gambar maupun video, sehingga manusia dapat memanfaatkannya dalam kehidupan sehari-hari seperti rekonstruksi adegan, sistem transportasi cerdas, deteksi sebuah peristiwa, deteksi objek, hingga pelacakan objek (You, Li, Tao, Ou, & Gong, 2014).

Selain *computer vision*, terdapat beberapa metode *deep learning* yang biasa digunakan dalam kecerdasan buatan. Salah satu metode *deep learning* yang banyak digunakan yaitu metode CNN. Pada penelitian mengenai pendeteksian pejalan kaki yang menggunakan metode CNN, peneliti menggunakan metode tersebut untuk mengotomatisasi proses ekstraksi fitur dari gambar yang terdapat pada CCTV. Peneliti tersebut mendapatkan hasil yang cukup memuaskan dengan nilai akurasi yang mencapai 71,13% (Acharya, Khoshelham, & Winter, 2017).

Terdapat pula metode RCNN, seperti yang pernah diteliti pada penelitian pendeteksian aksi yang berupa aksi duduk, menulis, membaca, dan lainnya. Dengan metode yang digunakan, peneliti mendapatkan hasil bahwa pendekatan yang

digunakan memiliki dasar yang sederhana dan kuat dalam mendeteksi sebuah aksi (J. Wu, Kuang, Wang, Zhang, & Wu, 2020).

Metode RCNN juga pernah dikembangkan dalam penelitian pendeteksian objek berbasis *region*. Penelitian ini menggunakan set data PASCAL VOC 2010 dan didapatkan mAP sebesar 53,7%, lebih besar dibandingkan dengan menggunakan pendekatan *spatial pyramid* dan *bag-of-visual-words* dengan mAP sebesar 35,1% (Girshick, Donahue, Darrell, & Malik, 2014). Selain itu, RCNN juga mengungguli metode OverFeat (Sermanet et al., 2014) pada set data yang besar seperti ILSVRC2013. RCNN memperoleh nilai mAP 31,4%, lebih besar 7,1% daripada menggunakan OverFeat dalam mendeteksi sebuah objek.

Pada penelitian yang dilakukan Girshick et al. dalam pendeteksian objek digunakan metode RCNN dan *selective search* sebagai pendeteksian proposal wilayah (Girshick et al., 2014). Namun menurut Li et al., metode *selective search* yang digunakan dinilai belum cukup baik untuk melakukan pendeteksian (Li et al., 2018). Metode proposal lainnya yang dapat digunakan yaitu *Region Proposal Network* (RPN). RPN diusulkan pertama kali pada penelitian dengan menggunakan *Faster RCNN* (Ren, He, Girshick, & Sun, 2017). Metode RPN dinilai mampu untuk mengekstraksi proposal yang lebih tepat, hal ini dikarenakan klasifikasi *foreground-background* dan regresi *bounding box* yang digunakan (Li et al., 2018).

Pada penelitian yang dilakukan oleh Yanagisawa et al., di mana peneliti tersebut melakukan penelitian mengenai deteksi objek yang dilakukan pada gambar manga dengan menggunakan metode RPN dan *selective search*. Hasil yang didapatkan pada penelitian yang dilakukan yaitu, metode RPN dapat mendeteksi gambar yang lebih ambigu pada gambar karakter maupun teks dibandingkan dengan *selective search* (Yanagisawa et al., 2018). Selain itu, terdapat penelitian pendeteksian teks dengan menggunakan RPN. Penelitian tersebut menggunakan dua tahap pendekatan yaitu proposal wilayah dengan metode RPN dan klasifikasi wilayah dengan RCNN. Pendekatan yang digunakan dapat mencapai hasil yang kompetitif pada set data ICDAR2015 dan ICDAR2013 (Jiang et al., 2017).

Penelitian deteksi objek pada video rekaman CCTV sudah pernah dilakukan sebelumnya, salah satunya yaitu pada pendeteksian manusia. Penelitian dilakukan dengan mengambil informasi dari karakteristik wanita dan pria yang terdapat pada

video pengawasan CCTV. Pada penelitian tersebut, digunakan metode *Faster RCNN* untuk mendeteksi. Setelah proses pendeteksian, dilakukan perbandingan menggunakan *Euclidean distance* dan Siamese. Hasil dari penelitian tersebut, didapati bahwa perhitungan yang dilakukan dengan *Euclidean distance* memberikan hasil yang menjanjikan sebagai metode asosiasi objek dalam menemukan kesamaan antara dua hal (Chahyati, Fanany, & Arymurthy, 2017).

Pada tahun 2016, terdapat penelitian terkait menggunakan metode Siamese untuk pelacakan objek (Bertinetto, Valmadre, Henriques, Vedaldi, & Torr, 2016). Pada penelitiannya, Bertinetto et al. menggunakan algoritma pelacakan dasar dengan jaringan Siamese konvolusi sepenuhnya (*fully convolutional*) pada set data ILSVRC15 untuk menangani pembelajaran matriks kesamaan deteksi objek dalam video. Menurutny, meskipun sederhana, algoritma yang diajukan Bertinetto et al. dapat bekerja secara *real-time*. Selain itu, Siamese juga diterapkan pada beberapa penelitian lainnya, diantaranya untuk mempelajari matriks kesamaan antara dua kalimat (Shi, Wang, & Sakai, 2020), dan melakukan verifikasi tanda tangan yang ditulis pada tablet menggunakan pena tablet (Bromley et al., 1993).

Namun, dalam pendeteksian objek pada video rekaman CCTV terdapat beberapa kesulitan. Dalam satu perangkat, CCTV dapat merekam banyak objek dalam suatu wilayah dari waktu ke waktu. Maka dari itu, dibutuhkan pendekatan yang tepat dalam mendeteksi objek pada video rekaman CCTV. Salah satu pendekatan yang dapat digunakan yaitu pendekatan objek referensi.

Pendekatan objek referensi merupakan pendeteksian objek berdasarkan referensi yang ada. Bagian gambar dari objek yang ingin dicari dari sebuah video akan diambil untuk dijadikan referensi atau pembanding dalam pendeteksian objek pada video rekaman CCTV. Berdasarkan penelitian yang telah dilakukan oleh Bertinetto et al. (Bertinetto et al., 2016), metode Siamese dapat digunakan untuk mempelajari matriks kesamaan. Oleh karena itu, metode Siamese tepat digunakan dalam pendekatan objek referensi ini.

Berdasarkan penelitian-penelitian yang sudah dilakukan sebelumnya, terdapat beberapa metode yang dapat digunakan dalam penelitian objek referensi pada video rekaman CCTV ini. Pada penelitian ini, akan menggabungkan metode RCNN dan Siamese untuk mendeteksi objek referensi, di mana proses pendeteksian

proposal wilayah akan diganti dengan metode RPN dan proses klasifikasi akan menggunakan metode Siamese untuk mencari kesamaan antar dua objek tersebut. Selain itu, penelitian akan dilakukan dengan membuat dua model berbeda dari penggabungan metode-metode tersebut, di mana salah satu dari model tersebut akan ditambahkan metode perhitungan *Euclidean distance*. Dengan penggabungan tersebut diharapkan mampu memberikan hasil deteksi objek referensi yang baik pada video rekaman CCTV.

1.2 Rumusan Masalah

Setelah melihat latar belakang yang ada, maka penulis merumuskan permasalahan yang akan diangkat dalam penelitian ini. Adapun rumusan masalah yang diambil adalah sebagai berikut:

1. Bagaimana penggabungan metode RCNN dengan Siamese untuk deteksi objek referensi pada video rekaman CCTV?
2. Bagaimana performa dari pendeteksian objek referensi dengan metode RCNN dengan Siamese pada video rekaman CCTV?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan penggabungan metode RCNN dengan Siamese untuk deteksi objek referensi pada video rekaman CCTV.
2. Menganalisis hasil performa dari pendeteksian objek referensi dengan menggabungkan metode RCNN dengan Siamese pada video rekaman CCTV.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Mendapatkan pengetahuan mengenai penggabungan metode RCNN dengan Siamese untuk objek referensi pada video rekaman CCTV.
2. Mendapatkan hasil performa dari pendeteksian objek referensi dengan menggabungkan metode RCNN dengan Siamese pada video rekaman CCTV.

1.5 Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut:

1. Pendeteksian objek referensi pada video rekaman CCTV tidak *real-time*.

2. Pengujian dilakukan menggunakan set data yang dibuat untuk penelitian ini.
3. Objek yang terdapat pada set data yang dibuat merupakan objek diam.

1.6 Sistematika

Sistematika dalam penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini mengandung latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

BAB II KAJIAN PUSTAKA

Pada bab ini akan dijelaskan teori yang mendukung penelitian. Teori yang dijelaskan yaitu mengenai deteksi objek, *deep learning*, metode RCNN, RPN serta jaringan Siamese.

BAB III METODOLOGI PENELITIAN

Bab ini berisi penjelasan langkah-langkah yang akan dilakukan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi penjelasan mengenai hasil penelitian dan eksperimen yang telah dilakukan.

BAB V KESIMPULAN

Bab ini berisi kesimpulan dari keseluruhan penelitian yang telah dilakukan serta saran dari penulis untuk kegiatan penelitian selanjutnya.